UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/808,326	03/25/2004	Shioupyn Shen	0026-0072	5423

44989 7590 08/19/2008

HARRITY SNYDER, LLP
11350 Random Hills Road
SUITE 600
FAIRFAX, VA 22030

EXAMINER	
SMITH, GARRETT A	

ART UNIT	PAPER NUMBER
2168	

MAIL DATE	DELIVERY MODE
08/19/2008	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/808,326	SHEN, SHIOUPYN	
	Examiner	Art Unit	
	Garrett Smith	2168	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *08 May 2008*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-34* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-34* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is regarding Applicant's response filed 8 May 2008 to a prior Office Action. Claims 1 – 34 are pending. Claims 1, 12, 20, 26, 31 and 34 are amended.

2. This Office Action is the **Third Action – Final Rejection**.

Response to Arguments

35 USC § 101

3. Applicant's arguments (page 12) and amendments, filed 8 May 2008, regarding the rejection under 35 USC § 101 of claims 20 – 30 and 34 have been fully considered and are persuasive. The Examiner notes that memory and a process appear to be only hardware and not software. For these reasons, the rejection under 35 USC § 101 of claim 20 – 30 and 34 is **withdrawn**.

35 USC § 102(a): Schleimer et al.

4. Applicant's arguments (page 13 – 14) and amendments, filed 8 May 2008, regarding the rejection under 35 USC § 102(a) of claims 1, 2, 5, 6, 8, 11, 13, 16, 17, 20, 24, 26, 27 and 28 have been fully considered and are persuasive-in-part.

Applicant argues that Schleimer et al. does not teach when choosing the subset of the plurality of overlapping blocks that the subset has less than all of the possible overlapping blocks. The Examiner agrees with Applicant. However, Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference

between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time).

As for Applicant's arguments regarding "overlapping blocks", the Examiner submits that in the generation of the example k-grams in Figure 1, the "document" in 1a is repeated to make the last k-gram a full k-gram. Thus, there are k-grams that are generated which can overlap a earlier section of the document. This is one of a number of possible ways to generate "overlap" in particular selected blocks. Applicant provides in the Specification a slightly different form of overlap by a single (or multiple) character shift (i.e, using the example document from Schleimer et al. "adorunrunrunadorunrun", the first block of four characters would be "ador", the second would "doru", the third "orun", and so forth with one character shift). While Applicant's claimed invention would cover both forms (and possibly a number of others), the Examiner would like to note that even if Applicant claimed specifically the character shift overlap method described in the Specification, this method would also be extremely obvious and well known by a person of ordinary skill in the art.

In regard to claim 2, Applicant argues, on page 17, that Schleimer et al does not disclose setting the bits of the fingerprint. Applicant further argues:

> Claim 2 does not simply recite that compacting the subset of the plurality of overlapping blocks includes setting bits. Claim 2 recites that compacting the subset of the plurality of overlapping blocks includes setting bits in the representation of the document based on the subset of the plurality of overlapping blocks.

While the Examiner agrees that this is recited in claim 2, the Examiner respectfully disagrees with Applicant's interpretation. The "based on" required that, in some way, the result of the step is affected by the item(s) or object(s) following the "based on" clause. The Examiner submits that storing in memory the result of fingerprint generation (which is based on the selected blocks) satisfies the requirements of the claim language.

In regard to claim 13, the Examiner submits that Schleimer et al. further discloses the selecting the predetermined number of "sample. On page 1, second column, first full paragraph of Schleimer et al., the reference discusses the selection of samples with the formula "0 mod p, for some fixed p". The modulus function allows for a predetermined number of sample (rather then all of the possible hashes). The remaining limitations are either equivalent to those presented in claim 1 or addressed in the discussion of claim 2.

For these reasons, the rejection under 35 USC § 102(a) of claims 1, 2, 5, 6, 8, 11, 20, 24, 26, 27 and 28 is **withdrawn**. The rejection under 35 USC § 102(a) of claims 13, 16 and 17 are **maintained**.

35 USC § 102(b): Burrows

5. Applicant's arguments (page 21 – 22) and amendments, filed 8 May 2008, regarding the rejection under 35 USC § 102(b) of claim 20 have been fully considered and they are persuasive in part.

Applicant argues, "Therefore, this figure of BURROWS cannot disclose or suggest a fingerprint creation component to generate a fingerprint of a predetermined length for an input document, the fingerprint generated by sampling the input document to obtain samples, choosing a subset of the samples, and generating the fingerprint from the subset of the samples by compacting the subset of the samples, as recited in claim 20." The Examiner respectfully disagrees. The parsing module 30, shown in Figure 2 and a result shown in Figure 5, takes the input page and "produces a sequence of pairs 500 in a collating order according to the locations of the words of the various pages" (col 9, lines 33 – 40). The result of the parsing module is then inputted into the indexing module (see Figure 2, element #400). The Examiner must note that the limitation "to obtain the representation of the document" is intended use. The parsed word representation of Figure 5 is all that is required to meet the claim limitation as it performs "compaction" when it generates the "sequence of pairs".

However, the Examiner notes that Burrows does not appear to teach the limitation "where the subset is less then the entirety of the plurality of overlapping blocks". Thus this argument is persuasive. However, as noted above, the Schleimer et al. reference selects a smaller subset (for example; four selected hashes in step Figure

1c) to compose the fingerprint. Thus, a rejection under 35 USC 103(a) is entered below for claim 20.

For these reasons, the rejection under 35 USC § 102(b) of claim 20 is **withdrawn**.

35 USC § 103(a): Burrows and Ward et al.

6. Applicant's arguments (page 23 - 28) and amendments, filed 8 May 2008, regarding the rejection under 35 USC § 103(a) of claims 1 – 2, 5 – 8 and 10 have been fully considered and are persuasive. For these reasons, the rejection under 35 USC § 103(a) of claims 1 – 2, 5 – 8 and 10 is **withdrawn**. However, as discussed above, a rejection under 35 USC 103(a) is entered below including Schleimer et al.

7. The Examiner further notes that Applicant challenges the Examiner's use of Official Notice (page 30 - 33). In this regard, the transversal is inadequate. MPEP 2144.03(c),

> To adequately traverse such a finding, an applicant must specifically point out the supposed errors in the examiner's action, which would include stating why the noticed fact is not considered to be common knowledge or well-known in the art. See 37 CFR 1.111(b).

The Examiner notes that Applicant has not specifically and clearly pointed out the deficiencies of the knowledge possessed by a person of ordinary skill in the art and why it would not be in the common body of knowledge. As such, the Official Notice stands.

Claim Rejections - 35 USC § 102(a)

8. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

9. Claims **13, 16 and 17** are rejected under 35 U.S.C. 102(a) as being anticipated by Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

10. In regard to **claim 13**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping samples (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3); and compacting the subset of the overlapping blocks includes setting bits in the representation of the document based on the subset of the overlapping blocks ((g) of Figure 2, bits of the representation of the document i.e. document signature are set).

11. In regard to **claim 16**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)) and choosing the smallest hash value in a window (Section 3).

12. In regard to **claim 17**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

Claim Rejections - 35 USC § 103

13. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

14. Claims **1, 2, 5, 6, 8, 11, 20, 24, 26, 27 and 28** are rejected under 35 U.S.C. 103(a) as being unpatentable over Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

15. In regard to **claim 1**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious

reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

16. In regard to **claim 2**, Schleimer et al discloses compacting the subset of the overlapping blocks includes setting bits in the representation of the document based on the subset of the overlapping blocks ((g) of Figure 2, bits of the representation of the document i.e. document signature are set).

17. In regard to **claim 5**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

18. In regard to **claim 6**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

19. In regard to **claim 8**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

20. In regard to **claim 11**, Schleimer et al discloses the overlapping blocks can be of the same length (Section 3).

21. In regard to **claim 20**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the

representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3). Schleimer et al further discloses comparing the document representation against a query based on the representation of another document (Section 3.2).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

22. In regard to **claim 24**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

23. In regard to **claim 26**, Schleimer et al discloses sampling the document to obtain a plurality of overlapping blocks (Section 3: Winnowing; windows can be overlapping sample of a document); choosing a subset of the overlapping blocks (Section 3: Winnowing; figure 2(e) shows a set of predetermined size of elements are selected into windows); and compacting the subset of the overlapping blocks to obtain the

representation of the document (figure 2(e) shows the compaction by winnowing, so does Section 3).

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

24. In regard to **claim 27**, Schleimer et al discloses hashing the overlapping blocks (figure 2(d)).

25. In regard to **claim 28**, Schleimer et al discloses choosing the smallest hash value in a window (Section 3).

26. Claims **1 – 2, 5 – 8 and 10** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and

Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003) and .

27. In regard to **claim 1**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have

been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

28. In regard to **claim 2**, Burrows teaches setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

29. In regard to **claim 5**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

30. In regard to **claim 6**, Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

31. In regard to **claim 7**, Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of

Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

32. In regard to **claim 8**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

33. In regard to **claim 10**, Ward et al teaches wherein setting the bits includes flipping a bit in the representation of the document when the bit corresponds to a block in the subset of plurality of overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks).

34. Claims **3, 4, 11 and 12** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows, Ward et al and Schleimer et al. as applied to claim 1 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

35. In regard to **claim 3**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the

method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

36. In regard to **claim 4**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

37. In regard to **claim 11**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a predetermined size such as 8 bytes (see col 6, lines 4 – 7). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

38. In regard to **claim 12**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al teaches that "words" can be of a

predetermined size such as 8 bytes (see col 6, lines 4 – 7) and under sized words can be padded to bring them to correct size. It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

39. Claims **13, 16 and 17** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

40. In regard to **claim 13**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

41. In regard to **claim 16**, Ward et al teaches hashing of the data blocks (see ¶40). Ward et al also teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40).It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

42. In regard to **claim 17**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

43. Claims **14 and 15** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows, Ward et al as applied to claim 13 above, and further in view of Broder et al (US Patent 6,230,155 B1; patented 8 May 2001).

44. In regard to **claim 14**, Burrows and Ward et al teach the invention as substantially claimed. Burrows and Ward et al do not explicitly teach that the representation of the document be of a predetermined length. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14). It would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents.

45. In regard to **claim 15**, Burrows and Ward et al teach the invention as substantially claimed. However, Broder et al does teach that a predetermined length of the representation of a document (see col 5, lines 1 – 14) and suggests that a longer fingerprint reduces the chance of two documents that are not similar that have exactly the same fingerprint (see col 5, lines 1 – 14). Therefore, it would have been obvious to a person of ordinary skill in the art at the time of invention to use the predetermined length of representation of a document of Broder et al with the method of Burrows and Ward et al because it allows for easy comparison of the fingerprints between two documents as well as reduce the chance of two documents that are not similar that have exactly the same fingerprint.

46. Claim **20** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

47. In regard to **claim 20**, Burrows discloses a fingerprint creation unit (2410) that samples the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); chooses a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows further discloses a similarity detection component to compare fingerprints to determine whether pairs of fingerprints correspond to near-duplicate documents (2420 and see figure 24).

However, the Examiner notes that Burrows does not appear to teach the limitation "where the subset is less then the entirety of the plurality of overlapping blocks". The Schleimer et al. reference selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. It would have been obvious to a person of ordinary skill in the art at the time of invention to use the selection capabilities of Schleimer et al. with the system/method of Burrows because it leads to improved efficiency.

48. Claims **22 and 23** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows and Schleimer et al. as applied to claim 20 above, and further in view of Charikar ("Similarity Estimation Techniques from Rounding Algorithms"; published 19 May 2002).

49. In regard to **claim 22**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

50. In regard to **claim 23**, Burrows teaches the invention as substantially claimed. Burrows does not explicitly state the use of Hamming Space for the comparison of

fingerprints. However, Charikar does teach use of Hamming space for calculating the similarity between fingerprints (see page 382, col 2, second paragraph). It would have been obvious to a person ordinary skill in the art at the time of invention to use the Hamming space calculations of Charikar with the components of Burrows because it is an able and suggested method for computing nearest neighbor problems and similarity tests.

51. Claims **24 and 25** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) and Schleimer et al. as applied to claim 20 above, and further in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

52. In regard to **claim 24**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

53. In regard to **claim 25**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

54. Claims **26 – 29** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) and Schleimer et al. in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002).

55. In regard to **claim 26**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); and compacting the subset of the overlapping blocks to obtain the representation of the document (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting

fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

56. In regard to **claim 27**, Burrows teaches the invention as substantially claimed. Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

57. In regard to **claim 28**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors. Another obvious choice inferred from Ward et al is the lowest weighted feature can be chosen (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

58. In regard to **claim 29**, Burrows teaches the invention as substantially claimed. Ward et al teaches choosing the highest weighted feature of the computed vectors (¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

59. Claim **31, 32 and 34** are rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al

(US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. ("Winnowing: Local Algorithms for Document Fingerprinting"; published 9 June 2003).

60. In regard to **claim 31**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have

been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

61. In regard to **claim 32**, Ward et al teaches hashing of the data blocks (see ¶40). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

62. In regard to **claim 34**, Burrows teaches sampling the document to obtain a plurality of blocks (see figure 4, blocks are selected from a document); choosing a subset of the overlapping blocks (see figure 4, blocks are selected from a document); setting bits in the representation of the document based on the subset of the overlapping blocks (Figure 5 shows the resultant compaction of the results of the selected overlapping blocks; since the representation is stored in memory or on a disk, bits are set based on the overlapping blocks). Burrows does not explicitly teach that the blocks can be overlap data. However, Ward et al teaches a sliding window of overlap for data (see ¶32). It would have been obvious to a person of ordinary skill in the art to use the window overlapping sampling of Ward et al with the method of Burrows because it allows for quicker indexing and a higher accuracy of the resulting samples.

Schleimer et al teaches the choosing of less then all of the k-gram hashes. Therefore, the difference between the claimed invention of claim 1 and the Schleimer et al. reference is that the claimed invention selects the subset of blocks and then hashes the selected blocks ("compacting") and the Schleimer et al. reference hashes all of the k-grams (selected blocks) and then selects a smaller subset (for example; four selected

hashes in step Figure 1c) to compose the fingerprint. Thus, this is an obvious reordering of steps that would be recognized by a person of ordinary skill in the art because it leads to increased efficiency and a reduction of data processing (selecting fewer then all blocks to hash reduces the amount of processor time). It would have been obvious to a person of ordinary skill in the art at the time of invention to reorder the steps of Schleimer et al. because it leads to improved efficiency.

63. Claim **9** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 8 above, and further in view of Official Notice.

64. In regard to **claim 9**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

65. Claim **18** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 16 above, and further in view of Official Notice.

66. In regard to **claim 18**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

67. Claim **19** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) as applied to claim 17, and further in view of Official Notice.

In regard to **claim 19**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

68. Claim **33** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB

2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 32 above, and further in view of Official Notice.

69. In regard to **claim 33**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that taking a number of least significant bits is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

70. Claim **21** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) as applied to claim 20 above, and further in view of Official Notice.

71. In regard to **claim 21**, Burrows also discloses a search engine (140, see figure 1). However, Burrows does not explicitly disclose returning a single link when the documents are determined to be duplicates. The Examiner takes Office Notice that returning a single link when the documents are determined to be duplicates is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use returning a single link with the components of Burrows because it would reduce the amount of data traffic and provide the user with clarity as to the nature of the document.

72. Claim **30** is rejected under 35 U.S.C. 103(a) as being unpatentable over Burrows (US Patent 5,745,900 B1; patented 28 April 1998) in view of Ward et al (US PGPUB 2002/0133499 A1; published 19 September 2002) and Schleimer et al. as applied to claim 27, and further in view of Official Notice.

73. In regard to **claim 30**, Burrows and Ward et al teach the invention as substantially claimed. The Examiner takes Official Notice that flipping bits based on a hash (as it done for generation of encryption keys via hashing) is well known by a person of ordinary skill in the art at the time of invention. It would have been obvious to a person of ordinary skill in the art at the time of invention to use this type of hashing technique in the method of Burrows and Ward et al because it is allows a convenient low overhead method a determining which bin a particular sample gets placed into.

Conclusion

74. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Garrett Smith whose telephone number is (571) 270-1764. The examiner can normally be reached on Mon - Fri, 8:30 AM - 6:00 PM EST, Alt Fri Off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tim T. Vo can be reached on (571) 272-3642. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

August 18, 2008

/GS/
Garrett Smith
Patent Examiner
Art Unit 2168

/Tim T. Vo/
Supervisory Patent Examiner, Art Unit 2168